

What is claimed is:

- 1 1. A method of discriminating speech from voice-band data in a
2 communication network, comprising:
3 calculating at least one of a self similarity ratio value, representing a
4 periodicity characteristic, and an autocorrelation coefficient value,
5 representing a spectral characteristic, for an input signal segment; and
6 determining whether said input signal segment is speech or voice-
7 band data based on said at least one of said self similarity value and said
8 autocorrelation coefficient value.
- 1 2. The invention as defined in claim 1, wherein said input signal
2 segment is a frame of N samples.
- 1 3. The invention as defined in claim 1, wherein
2 said calculating step calculates a first self similarity ratio value,
3 corresponding to a first sample delay, as a first periodicity characteristic
4 value; and
5 said determining step determines that said input signal segment is
6 voice-band data if said first self similarity ratio value is greater than a first
7 similarity threshold.
- 1 4. The invention as defined in claim 3, wherein
2 said calculating step calculates a second self similarity ratio value,
3 corresponding to a second sample delay, as a second periodicity
4 characteristic value, said second sample delay being greater than said first
5 sample delay; and
6 said determining step determines that said input signal segment is
7 speech if said second self similarity ratio value is greater than a second
8 similarity threshold.

1 5. The invention as defined in 1, wherein

2 said calculating step calculates a first autocorrelation coefficient as a
3 first spectral characteristic value; and

4 said determining step determines that said input signal segment is
5 voice-band data if said first autocorrelation coefficient is less than a first
6 autocorrelation threshold, and that said input signal segment is speech if
7 said first autocorrelation coefficient is greater than a second
8 autocorrelation threshold, said second autocorrelation threshold being
9 greater than said first autocorrelation threshold.

1 6. The invention as defined in claim 5, wherein

2 said calculating step calculates second and third autocorrelation
3 coefficients as second and third spectral characteristic values respectively,
4 and

5 said determining step determines that said input signal segment is
6 voice-band data if said second autocorrelation coefficient is less than a
7 third autocorrelation threshold or said third autocorrelation coefficient is
8 less than a fourth autocorrelation threshold.

1 7. The invention as defined in claim 6, wherein

2 said determining step determines that said input signal segment is
3 voice-band data if a sum of said second autocorrelation coefficient and said
4 third autocorrelation coefficient is less than a fifth autocorrelation
5 threshold.

1 8. The invention as defined in claim 1, wherein

2 said calculating and determining steps are performed for a plurality
3 of input signal segments in accordance with a sequential decision logic
4 sequence which designates input signal segments as speech during a

speech state and designates input signal segments as voice-band data during a voice-band data state.

9. The invention as defined in claim 8, wherein

said sequential decision logic sequence switches from said speech state to said voice-band data state when results of said determining step for a plurality of input signal segments indicate that said speech state is erroneous, and

said sequential decision logic sequence switches from said voice-band data state to said speech state when results of said determining step for a plurality of input signal segments indicate that said voice-band data state is erroneous.

10. The invention as defined in claim 8, wherein

results of said determining step are weighted based on energy content of the corresponding input signal segment so that determination results for low energy input signal segments are given relatively low weight when determining whether to switch from said speech state to said voice-band data state or from said voice-band data state to said speech state.

11. An apparatus for discriminating speech from voice-band data in a communication network, comprising:

calculating means for calculating at least one of a self similarity ratio value, representing a periodicity characteristic, and an autocorrelation coefficient value, representing a spectral characteristic, for an input signal segment; and

determining means for determining whether said input signal segment is speech or voice-band data based on said at least one of said self similarity value and said autocorrelation coefficient value.

1 12. The invention as defined in claim 11, wherein said input signal
2 segment is a frame of N samples.

1 13. The invention as defined in claim 11, wherein
2 said calculating means calculates a first self similarity ratio value,
3 corresponding to a first sample delay, as a first periodicity characteristic
4 value; and

5 said determining means determines that said input signal segment
6 is voice-band data if said first self similarity ratio value is greater than a
7 first similarity threshold.

1 14. The invention as defined in claim 13, wherein
2 said calculating means calculates a second self similarity ratio value,
3 corresponding to a second sample delay, as a second periodicity
4 characteristic value, said second sample delay being greater than said first
5 sample delay; and

6 said determining means determines that said input signal segment
7 is speech if said second self similarity ratio value is greater than a second
8 similarity threshold.

1 15. The invention as defined in 11, wherein
2 said calculating means calculates a first autocorrelation coefficient
3 as a first spectral characteristic value; and

4 said determining means determines that said input signal segment
5 is voice-band data if said first autocorrelation coefficient is less than a first
6 autocorrelation threshold, and that said input signal segment is speech if
7 said first autocorrelation coefficient is greater than a second
8 autocorrelation threshold, said second autocorrelation threshold being
9 greater than said first autocorrelation threshold.

1 16. The invention as defined in claim 15, wherein
2 said calculating means calculates second and third autocorrelation
3 coefficients as second and third spectral characteristic values respectively,
4 and
5 said determining means determines that said input signal segment
6 is voice-band data if said second autocorrelation coefficient is less than a
7 third autocorrelation threshold or said third autocorrelation coefficient is
8 less than a fourth autocorrelation threshold.

1 17. The invention as defined in claim 16, wherein
2 said determining means determines that said input signal segment
3 is voice-band data if a sum of said second autocorrelation coefficient and
4 said third autocorrelation coefficient is less than a fifth autocorrelation
5 threshold.

1 18. The invention as defined in claim 11, wherein
2 said apparatus classifies a plurality of input signal segments as
3 being either speech or voice-band data in accordance with a sequential
4 decision logic sequence which designates input signal segments as speech
5 during a speech state and designates input signal segments as voice-band
6 data during a voice-band data state.

1 19. The invention as defined in claim 18, wherein
2 said apparatus, in accordance with said sequential decision logic
3 sequence, switches from said speech state to said voice-band data state
4 when results of said determining means for a plurality of input signal
5 segments indicate that said speech state is erroneous, and
6 said apparatus, in accordance with said sequential decision logic
7 sequence, switches from said voice-band data state to said speech state

8 when results of said determining means for a plurality of input signal
9 segments indicate that said voice-band state is erroneous.

1 20. The invention as defined in claim 18, wherein
2 said apparatus weights results of said determining means based on
3 energy content of the corresponding input signal segment so that
4 determination results for low energy input signal segments are given
5 relatively low weight when said apparatus judges whether to switch from
6 said speech state to said voice-band data state or from said voice-band
7 data state to said speech state.